Does grade retention invoke student misbehavior in adolescence? A multilevel approach.

Paper presented at ECER, 18th – 21th September 2012, in Cadiz, Spain.
Forthcoming in 'School Effectiveness and School Improvement'

Jannick Demanet[a,1] and Mieke Van Houtte[a]

[a]Research group CuDOS, Department of Sociology, Ghent University, Gent, Belgium
[1]Corresponding author. Korte Meer 5, 9000 Gent, Belgium. Tel.: +32 (0)9/264.91.90. Mail: Jannick.Demanet@ugent.be

Abstract

Few studies with respect to grade retention and school-disruptive behavior have focused on adolescence. Moreover, previous retention research has ignored multilevel issues. This study aims to fill these research lacunae by addressing the role of grade retention in adolescent students' school misconduct. Furthermore, we explore the role of the percentage of retained students at school in individual-level school misconduct and in moderating the relationship between retention and misconduct. Multilevel analyses of data (2004-2005) from 11,872 students in 85 Flemish secondary schools suggest that, while students retained in primary education exhibited less school misconduct in adolescence, those retained in secondary schools were more likely to break rules. Furthermore, students attending schools with a higher percentage of retainees were found to be more deviant. However, schools' retention composition moderated the relationship between grade retention and school misconduct. Implications are discussed.

INTRODUCTION

In many countries, the practice of grade retention is widespread (*Switzerland:* Bonvin, Bless, & Schuepbach, 2008; *Germany:* Ehmke, Drechsel, & Carstensen, 2010; *US:* Jimerson, 2001; Lorence & Dworkin, 2006; *Canada:* Pagani, Tremblay, Vitaro, Boulerice, & McDuff, 2001; *Belgium:* Juchtmans et al., 2011; Van Petegem & Schuermans, 2005). Proponents believe that giving students "the gift of time" will put them back on track for normal educational growth. Ensuing the popularity of this strategy, a rich body of research has developed to test its effectiveness. The practice has some positive effects on students' cognitive growth (Alexander, Entwisle, & Dauber, 1994), but these remain mainly short-term (Jimerson & Ferguson, 2007; Meisels & Liaw, 1993) and occur only when special help is provided to retained children. In fact, in recent decades many studies have condemned grade retention as an ineffective practice to improve student learning (see e.g., Bonvin et al., 2008; Jimerson, 2001; Mccoy & Reynolds, 1999; Pagani et al., 2001).

Although most studies on grade retention have pinpointed cognitive outcomes, others have linked retention to a range of problematic behavioral outcomes. Most studies in this area have linked grade retention to school dropout (e.g., Jimerson, Anderson, & Whipple, 2002; Stearns, Moller, Blau, & Potochnick, 2007), consistently finding that retained students have a higher chance of dropping out of school. Fewer have focused on students' behavior at school (Gottfredson, Fink, & Graham, 1994; Jimerson & Ferguson, 2007). A shortcoming of the latter, however, is that most studies on school misbehavior have focused on primary school (see, e.g., Mccoy & Reynolds, 1999; Pagani et al., 2001) or middle school contexts (e.g., Gottfredson et al., 1994), while retention effects on deviancy during adolescence have been virtually ignored (for a notable exception, see Jimerson & Ferguson, 2007). However, researchers have stated that the outcomes of retention may be different for students in different life phases, as older retainees would feel more stigmatized than younger ones (Finlayson, 1977; Wu, West, & Hughes, 2010). We would expect, consequently, that retention may yield school deviancy, especially in adolescent years. The first aim of this study is to fill the gap in the scientific literature by focusing on the relationship between grade retention and misbehavior in adolescent students.

Scholars have pointed out that retention research is methodologically flawed in a number of ways (see, e.g., Lorence, 2006). Most notably, it has failed to account for the multilevel nature of the school context (see Hong & Raudenbush, 2005), since very few studies have investigated school-level influences (Hong & Raudenbush, 2006). Previous research has shown that schools differ greatly in their retention policies, which yield differences in the schools' retention composition. Hong and Raudenbush (2005; 2006) assessed the effect of this compositional characteristic on students' performance. However, in the area of school deviancy, no study has investigated yet possible effects of the percentage of students retained. These authors also hold that previous retention research has started from the stable unit treatment value assumption (SUTVA; Rubin, 1986), meaning that the individual outcome of an intervention strategy is independent of the treatment other individuals receive. However, as students interact with each other at school, that assumption is untenable in a multilevel school context and, consequently, it is possible that the effects of being retained depend on the percentage of retained students in school (Hong & Raudenbush, 2006). While these multilevel issues have been investigated in the context of students' cognitive outcomes (Hong & Raudenbush, 2005; Hong & Raudenbush, 2006), they are still unexplored with regard to behavioral outcomes (Hong & Yu, 2008).

In short, this study addresses three research questions: first, whether there is a relation between retention and misconduct in adolescence, second, what is the effect of the percentage of retained students at school on the development of school misconduct, and, third, whether an association between retention and school misconduct is dependent upon the percentage of students retained in school.

Background

Grade retention research and its methodological flaws

In most educational systems, schooling is organized by assigning children to homogeneously organized age groups. Children make transitions through age-based grade levels (Hong & Raudenbush, 2005). This is intended to render grades as homogeneous as

possible with regard to students' academic ability. However not all students progress at the same academic pace, and while grades are homogeneous with regard to age, they are not always so with regard to academic ability (Ehmke et al., 2010). One possible way to deal with this increasing heterogeneity is grade retention. The practice of denying academically challenged students access to the next grade is intended to grant them the opportunity to review the material they failed to master during the previous year and eventually to catch up to normal educational growth.

Controversial from the start, the practice of grade retention has sparked researchers' interest in its effectiveness (for reviews, see Holmes, 1989; Jackson, 1975; Jimerson, 2001; Lorence, 2006). In two influential meta-analyses, Holmes (1989), reviewing 63 studies from 1960 to 1987, and Jimerson (2001), drawing on 22 studies largely performed during the 1990s, both conclude that grade retention is ineffective as a remedy for poor academic progress. Some studies have found a positive learning effect in primary schools (e.g. Alexander et al., 1994), but these positive effects only surface when retainees are provided with special help (Peterson, Degracie, & Ayabe, 1987). Moreover, when positive learning effects are found, they mostly remain short-term. For instance, researchers pointed to a "grade-replacement effect" (see Jimerson & Ferguson, 2007; Meisels & Liaw, 1993): while retained students perform better in the repeated grade, this beneficial effect is substantially reduced when they pass to a new grade (Alexander et al., 1994; Jimerson, Carlson, Rotert, Egeland, & Sroufe, 1997; Shepard & Smith, 1990). Hence the temporary improvement would simply be due to repetition of the course material. Consequently, most studies find that grade retention has no lasting beneficial effect on educational attainment (Jimerson, 2001; Shepard & Smith, 1990).

In recent years, research on retention has surfaced that criticizes earlier conclusions on the basis of methodological flaws (e.g., Ehmke et al., 2010; Lorence, 2006; Lorence & Dworkin, 2006). It points out that they have mainly been based on small sample sizes and unrepresentative data (Ehmke et al., 2010). In fact most sample sizes are so small that Lorence & Dworkin (2006) have stated that studies have lacked the statistical power to test retention's effectiveness. A second important shortcoming of retention research is that very few studies have accounted for the multilevel nature of the school context (for notable exceptions, see Hong & Raudenbush, 2005; Hong & Raudenbush, 2006). However, it has been demonstrated that there are differences between schools' retention policies, which yield differences in their retention composition (Shepard & Smith, 1988). Hence it is important to investigate the compositional effect of the percentage of retained students, which can possibly affect both retained and promoted students (Hong & Raudenbush, 2006). Furthermore, in the multilevel context of a schooling environment, one cannot assume that these differential compositions have no impact on the consequences of being retained as an individual student. In other words, the stable unit treatment value assumption (SUTVA, see Rubin, 1986), which holds that the outcome of an intervention strategy for a student is independent of the treatment other students receive, is untenable in the educational context (see Hong & Raudenbush, 2005; Hong & Raudenbush, 2006). These considerations led Hong & Raudenbush (2005; 2006) to assess the impact of individual retention on academic achievement both in low-retention and high-retention schools, and to assess the impact of those schools' retention composition on the achievement of promoted students. They found that in both high- and low-

retention schools, retained students' educational progress was less than that of children who were promoted (Hong & Raudenbush, 2006). Furthermore, they found no effect of that composition on promoted students. The authors interpreted this as evidence that retention is not an effective way of improving performance, either for retained students or for those promoted. According to the authors, however, the failure to find a compositional effect on academic achievement might be due to the study's crude distinction between low- and high-retention schools (Hong & Raudenbush, 2006, p. 909). Furthermore, there are theoretical reasons to expect that, in the case of school misconduct, the percentage of retained students does make a difference in the effect of grade retention for both those retained and promoted.

Grade retention and school misconduct

A number of studies have focused on retained students' likelihood of school-disruptive behavior (Gottfredson et al., 1994; Jimerson & Ferguson, 2007; Pagani et al., 2001). Most studies hold that the frustration of being retained yields a higher incidence of aggression and oppositional behavior (Jimerson & Ferguson, 2007; Pagani et al., 2001). Pagani and colleagues (2001) distinguished between short- and long-term effects, finding that retention yielded a short-term increase in school-disruptive behavior for girls and boys. It also resulted in dramatic and long-lasting effects on disruptive behavior for boys, especially when retention occurred early in their school career. However, retention research on deviancy outcomes provides mixed results, with other studies finding no association between retention and later disruptive behavior (Mccoy & Reynolds, 1999), or finding that retention diminishes students' deviancy (Gottfredson et al., 1994). According to Wu and colleagues (2010), these differences may, among other factors, be due to the timing of retention in the observed samples.

It is noteworthy that most studies on behavioral outcomes have been conducted in primary schools (e.g., Pagani et al., 2001) or middle schools (e.g., Gottfredson et al., 1994), while few studies have focused on adolescence (Jimerson & Ferguson, 2007). However, it is especially in adolescence that we may expect grade retention to lead to school misconduct. In past research, scholars have used social comparison theory (Festinger, 1954) to explain the short-term beneficial outcomes of grade retention in primary education (see Hong & Yu, 2008; Wu et al., 2010). Social comparison is the process by which individuals use others as a yardstick to evaluate their own situation. In the context of adolescence, however, we may expect these social comparison processes to produce more school misconduct among retained students. Researchers have suggested that older students may feel more stigmatized by retention than younger ones (Finlayson, 1977; Wu et al., 2010, p. 148). Consequently, we may expect that social comparison for adolescent retainees results more readily in relative deprivation. Since relative deprivation promotes adolescent deviancy (Rosenberg, Schooler, & Schoenbach, 1989), the experience of retention might yield school-disruptive behavior in adolescence. Supporting this argument, Jimerson and Ferguson (2007), in one of the few studies on adolescent outcomes, found that being retained increases the chances for deviancy in adolescent years. Given the dearth of research in this developmental period, it is important to perform more research on deviancy outcomes in adolescence.

Furthermore, while some studies have investigated students' cognitive outcomes by multilevel approaches (Hong & Raudenbush, 2005; Hong & Raudenbush, 2006), none have

applied multilevel procedures to the topic of student deviance. However it is important to account for school-level effects in at least two ways. First, no study has yet investigated the effects of the percentage of retained students in school on individual students' likelihood of school misconduct. Second, especially in the case of retention's effects on adolescent deviancy, the stable unit treatment value assumption (Rubin, 1986) cannot be maintained. Previous research has hinted that structural school characteristics, such as the tracking system (Van Houtte, Demanet, & Stevens, 2012), and compositional characteristics, such as the SES and ethnic composition (Demanet & Van Houtte, 2011), may shape students' frame of reference for making social comparisons. Similarly, we may expect the percentage of retained students at a school to shape the nature of the social comparison process, which may affect the relation between grade retention and school misconduct. Richer (1976) believes that feelings of relative deprivation depend upon social interaction. In other words, the visibility of the other group is important in social comparison processes. As retained students in low-retention schools may be expected to compare their own situation more readily to never-retained students than retained students may be in high-retention schools, social comparison in low-retention schools might be expected to result in greater relative deprivation than in high-retention schools. If so, then we can expect that school misconduct is more likely to be a result of retention in low-retention schools than in high-retention schools. This could explain why Gottfredson and colleagues (1994), contrary to most other studies, found a small deviance-preventing effect of grade retention: students retained in their sample had a lower incidence of deviant behavior. However, the schools investigated had a very high number of retained children. It might be, then, that processes of social comparison turned out less unfavorable for retained students, making deviancy less likely to occur. This possibility was forwarded by Gottfredson and colleagues (1994), but was dismissed because of the small deviance-preventing effect of retention they found. However, the theoretical proposition that retention is more deviance-yielding in low-retention schools than in high-retention ones has never been investigated empirically.

This literature overview leads us to address three research questions:

- Is being retained associated with school misconduct in adolescence?
- Is the percentage of retained students in a school related to school misconduct?
- Is the association between grade retention and school misconduct stronger in low-retention schools than in high-retention schools?

The Flemish educational context

Before we explain our methodological framework, a word is in order about the educational system in the Flemish context—Flanders is the Dutch-speaking, northern part of Belgium. Since 1988, the Flemish government has had the jurisdiction to implement and govern its own educational system, which limits the study to the students and schools in this region. First it should be kept in mind that every school in Flanders is state subsidized – public and private schools alike. With only a few exceptions, private schools are mainly Catholic. Public schools are non-sectarian. Usually children go to nursery school from the age of two and a half. Education becomes compulsory when the child is six years old. After six grades of primary education, at the age of twelve, children transfer to secondary education.

There are six grades of secondary education divided into three units, subdivided into two grades each. There are four main tracks in secondary education: academic education preparing for higher education; technical education; vocational education and artistic education (which is marginal in terms of number of students). Tracks are not only organized within but also, and mainly, between schools. The Flemish school system can be categorized as "explicit school-level tracking to different school types catering to specific student groups", using achievement as a selection criterion (Trautwein, Ludtke, Marsh, Koller, & Baumert, 2006, p. 789). The different tracks are commonly classified hierarchically, placing vocational tracks at the lower end.

At the end of each year students receive a certificate indicating whether they can continue their current school career (certificate A), or not (certificate B or C). A certificate B indicates that the student may pass to the next grade but needs to join a lower track; a certificate C means that the student is to be retained in the current grade. These certificates are based on the GPA obtained. It is important to note that, in the Flemish educational system, there are no standardized tests (for example in the form of centrally administered and standardized examinations), which makes educational achievement very hard to compare across schools and across students (Stevens, 2007). A "cascade-effect" has originated, in which students start in higher tracks, but when they fail to gain the necessary academic credentials, they move to lower tracks. Students in vocational tracks are more likely to have been retained in the past (Juchtmans et al., 2011). Hence attending a vocational track in Flanders is rarely a positive choice, and vocational students are all too aware of their low status in society, leading to more anti-school attitudes and misconduct (for an extended discussion of the Flemish tracking system, see Van Houtte et al., 2012; Van Houtte & Stevens, 2008).

METHODS

Research Design

To answer our research questions we used multilevel modeling (HLM6; Raudenbush & Bryk, 2002). As is common for delinquency measures (Crosnoe, Erickson, & Dornbusch, 2002; Stewart, 2003), the dependent variable was significantly skewed (see below) toward its lower end, violating the normality assumption underlying linear models. It was therefore more appropriate to treat the variable as an event or count variable. Hence, we estimated the effects using an overdispersed Poisson model with constant exposure (Long, 1997; Raudenbush & Bryk, 2002).

It is common in multilevel analyses to start by estimating unconditional models—that is without specifying any determinant—to determine the amount of variance that occurs among schools, but in hierarchical non-linear models it is not appropriate to partition the variance in the outcome into its between and within components (Frost, 2007; Lee & Burkam, 2003; Stewart, 2003). The between-school variance component τ_0 estimated in an unconditional model does give an idea of whether or not the between-school variance is significant and can be modeled. The results of the unconditional model showed that the level of school misconduct varied significantly across schools (τ_0=0.006; p<0.001), so that multilevel analysis was appropriate. To determine possible mediation effects we estimated stepwise multilevel regression models. Throughout all analyses, we controlled for several variables at

the school and individual level. At the school level, we controlled for the school's sector (coding: 1=public sector), SES composition, ethnic composition, and size, as these variables have been shown to affect school misconduct (Demanet & Van Houtte, 2011; Stewart, 2003; Stretesky & Hogan, 2005). We should point out that a high correlation existed between SES and ethnic school composition (r=-0.777; p<0.01). Given that other studies have established the effect of ethnic composition on school deviancy in addition to the effect of SES composition (e.g., Demanet & Van Houtte, 2011; Eitle & Eitle, 2003), we chose to use these variables simultaneously, although the results should be regarded with caution due to possible multicollinearity.

At the individual level, we controlled for the sociodemographic characteristics of gender (coding: 1=girl), age, ethnicity (coding: 1=ethnic minority student), and SES (Barone et al., 1995; Demanet & Van Houtte, 2011; Tygart, 1988). It was especially important to account for the respondents' age, as retained students are older than their classmates, and we had to be careful not to confound retention effects with age effects (see also Byrd, Weitzman, & Auinger, 1997). Additionally, we controlled for whether students attended the vocational track (Van Houtte & Stevens, 2008, see also "context"-section; coding: 1=vocational track). Lastly, we controlled for students' prior achievement, as students with lower prior achievement are more likely both to have been retained in the past, and to misbehave at school (Jimerson & Ferguson, 2007; Mccoy & Reynolds, 1999; Roeser & Eccles, 1998).

Because schooling is organized into two different systems in Flemish education – a primary school system and a secondary school system - and previous researchers have suggested that the timing of grade retention may matter in producing school deviancy (Wu et al., 2010), we distinguished between retention in primary and secondary education in our analyses. In the first step of the multilevel analyses, we investigated the role of retention in primary and secondary schools in students' school misconduct in secondary school. In the second step, we added the school percentage of students retained in secondary school to the model. A compositional effect arises when the composite effect is found to be significant over and above the individual effect. In the third step, we estimated a cross-level interaction effect between the proportion of retained students in school and retention in secondary education. This enabled us to test whether an eventual effect of grade retention on school misconduct varied by the school's retention composition. To ensure model stability, all independents but the dichotomous variables were grand mean centered. Because of the lack of standardized test scores in Flanders, it is however likely that there is considerable between-school variance in how students are evaluated. For this reason, in our analyses, the variable representing prior achievement was group mean centered. Of relevance here is whether or not the relative position of the respondent (that is the student) in the group (that is the school) affected the dependent variables (Opdenakker & Van Damme, 1997).

Procedure

We used retrospective data, gathered from students in secondary schools. The data were part of the FlEA (Flemish Educational Assessment), gathered in the 2004–2005 school year in 85 Flemish secondary schools. We used multistage sampling. First we selected proportional-to-size postal codes, size being defined by the number of schools within each postal code, information provided by the Educational Department. From the 240 postal codes, we selected

48 at random. This resulted in the desired overrepresentation of larger municipalities. Consequently, we selected all regular secondary schools in the chosen postal codes that provided a third and fifth grade (corresponding to years 9 and 11 in the US system), yielding a response rate of 31%. This low response rate is due to schools in Flanders being swamped with research requests. Schools choose the research they take part in on a first-come, first-served basis. Analyses in which we compared our sample to the Flemish school population, based on information attained through the Flemish Educational Department, showed that the participating schools did not differ from those that opted out in terms of school sector, size, curriculum, or student composition. Hence no systematic biases occurred, and the 85 schools in the sample are representative of the Flemish situation (Van Houtte et al., 2005).

In the participating schools, we asked all third and fifth grade students present at the time of the visit to fill out the questionnaire. Students filled out the questionnaire in class, supervised by members of the research team and a teacher. A few students were not present, due to absence or field trips. A total of 11,945 students completed the questionnaire, of which 11,872 (response rate: 87%) proved valid: 6,081 (response rate: 90%) in the third grade, 5,791 (response rate: 86%) in the fifth grade. The questionnaires were not anonymous because we needed to couple other data provided by the school with the students' responses. Ultimately we removed all names, so that all analyses were performed on anonymous data.

Sample characteristics

Our final sample consisted of 11,872 students across 85 schools. The sample was quite equally divided by gender (51.4% girls). Of the respondents, 51.20% attended the third grade. Hence the majority of students are 15 (34.8%) or 17 (32.6%) years old, with other students a little older than most in their grade, mainly due to grade retention (11.3% being 16 years old; 14.3% 18 years old; 4.6% 19 years old, and 1.4% being 20 years old). The mean age in the sample was 16.45 (SD=1.30; see Table 1). The majority of respondents were ethnic majority students (88.8%), which we defined as Western-Europeans, from which all grandparents are born in Western-Europe (e.g., Timmerman et al., 2002). Most members of ethnic minority groups had Turkish or Moroccan backgrounds (both about 30%), some had Southern-European (10%), Eastern-European (8%), North-African (5%), or other (16%) backgrounds. Respondents originated from families that covered the entire range of SES (1=unskilled manual labor; 8=professionals and large proprietors; see Erikson et al., 1979). The mean SES was 5.20 (SD=2.10; see Table 1). Most respondents attended the general track (46.7%), with 28.5% attending the technical, 22.1% the vocational, and 2.7% the arts track. To grade their students, Flemish high schools use a percentage, hence, grades range from 0% to 100%. In our sample, students' GPA (Grade Point Average) ranged from 41% to 100%, with a mean of 69.42% (SD=9.22; see Table 1).

Schools were equally divided across school sector. In our sample, 50.6% of the schools were public, which is a slight overrepresentation of the Flemish situation. This is because we oversampled larger municipalities, where the majority of public schools in Flanders are situated. Schools had an average SES composition of 4.80 (SD=1.23). The 85 schools in our sample covered the entire range of ethnic composition, from 0% (6 schools) to 88.20% (1 school) ethnic minority students. The mean ethnic composition was 16.45 (SD=21.70). Schools in our sample had an average size of 461.55 students (SD=285.27). However, we

obtained information on school size from only 83 of the 85 schools. As multilevel analysis does not permit missing values at the second level, and our analyses showed that school size exerts no influence on school misconduct, we eventually omitted this variable from all analyses.

Table 1: Sample characteristics: Frequencies (%), means (M), standard deviations (SD), Cronbach's alpha, and *N*.

Variables		%	M	SD	Cronbach's alpha	*N*
Dependent						
School misconduct			30.04	8.47	0.87	11,561
School level						
School sector						
	Public	50.60%				85
SES composition			4.80	1.23		85
Ethnic composition			16.45	21.70		85
School size			461.55	285.27		83
Percentage of retained students			22.61	14.87		85
Student level						
Gender						
	Girls	51.40%				11,843
Age			16.45	1.30		11,803
Ethnicity						
	Immigrant	11.20%				11,870
Vocational track						
	Vocational	22.10%				11,872
SES			5.20	2.10		11,137
Prior achievement			69.42	9.22		10,713
Retention primary						
	Retained	15.10%				11,744
Retention secondary						
	Retained	20.30%				11,543

15.10% of our respondents indicated that they had been retained at least once in the course of primary education (see Table 1). A total of 20.30% of the respondents in the sample had been retained at least once in secondary education. There was a very weak association

between grade retention in primary education and retention in secondary education (Cramer's V=0.027; $p<0.01$). Of the students retained in secondary education, 16.42% were retained in primary education as well. Of all respondents, 3.33% were retained both in primary and secondary education. There was also a substantial variation between schools in the percentage of retained students at school. School in our sample ranged from 0% (1 school) to 58% (1 school) retained students. On average, schools enrolled 22.61% (SD=14.87; see Table 1) students who had been retained at least once during the course of secondary education.

Dependent variable

We measured school misconduct using a scale inspired by Stewart (2003, p. 602-604), consisting of 17 items. The scale measures behavior that counters the school rules, even if it is relatively harmless, such as cheating on tests, skipping school, vandalism at school, and so forth (see appendix for all items). Students were asked how often they performed these deviant acts. Students could answer using a 5-point scale, ranging from *never* (score 1) to *very often* (score 5). A confirmatory factor analysis (PCA, extracting one component) confirmed the high loadings of these 17 items on one component (ranging from 0.457 to 0.736). Moreover, a Cronbach's alpha of 0.87 (N=11,561) confirmed the reliability of the scale. Scores were summed to a scale ranging from 17 to 85 (mean=30.04, SD=8.47, see Table 1). It has been shown that using self-reported measures is not ideal for measuring deviant acts (Crosnoe, 2002). However, it nonetheless remains the most common method of gathering such information (e.g., Stewart, 2003). We interpolated missing values by item correlation substitution (Huisman, 1999) so that a missing item was assigned the value of the most highly correlated item. As is common for delinquency measures (Stewart, 2003), the dependent variable was significantly skewed (1.58, SE=0.023) toward its lower end.

RESULTS

Grade retention and school misconduct

The answer to the first research question, whether grade retention is associated with school misconduct, is provided in model 1 (see Table 2). Retention in both primary and secondary school was related to misconduct, but while having been retained in primary education had a negative effect on school misconduct in secondary education (γ=-0.0324; $p<0.001$), retention in secondary education had a positive effect (γ=0.0804; $p<0.001$). All else being equal, whereas the misconduct score for a student who was promoted throughout primary education was 32.2817 (=Exp(3.4745); see Long, 1997), it was 31.2525 for a student retained during primary education (=Exp(3.4745-0.0324)). Conversely, all else being equal, having been retained in secondary education increased the expected school misconduct rate by 1.0837 (=Exp(0.0804)).

Although not the focus of the current study, we should point out that some other variables had effects on school misconduct. Students in public schools, male students, older students, ethnic majority students, students in the vocational track, those with a higher SES, and students with lower prior achievement, were all found to have a higher likelihood of being deviant at school (see Table 2). These results correspond with previous research (Barone et al., 1995; Demanet & Van Houtte, 2012; Roeser & Eccles, 1998; Tygart, 1988; Van Houtte & Stevens, 2008).

Table 2: Association between primary school retention, secondary school retention, percentage of retained students, and school misconduct. Results of stepwise multilevel analysis.

Variables		**Model 1**	**Model 2**	**Model 3**
Intercept		3.4745***	3.4833***	3.4837***
		(0.0158)	(0.0152)	(0.0153)
School level				
	School sector	0.0591***	0.0431***	0.0442***
		(0.0106)	(0.0102)	(0.0103)
	SES composition	-0.0095	0.0018	0.0017
		(0.0067)	(0.0073)	(0.0073)
	Ethnic composition	-0.0002	-0.0002	-0.0002
		(0.0004)	(0.0004)	(0.0004)
	Percentage of retained students		0.0018***	0.0019***
			(0.0005)	(0.0005)
Student level				
	Gender	-0.0811***	-0.0822***	-0.0821***
		(0.0079)	(0.0077)	(0.0076)
	Age	0.0229***	0.0223***	0.0223***
		(0.0035)	(0.0035)	(0.0035)
	Ethnicity	-0.0449**	-0.0441**	-0.0437**
		(0.0141)	(0.0142)	(0.0141)
	SES	0.0048**	0.0049**	0.0050**
		(0.0016)	(0.0016)	(0.0016)
	Vocational track	0.0544***	0.0526***	0.0514***
		(0.0111)	(0.0109)	(0.0110)
	Prior achievement	-0.0044***	-0.0044***	-0.0045***
		(0.0004)	(0.0004)	(0.0004)
	Retention primary	-0.0324***	-0.0319***	-0.0323***
		(0.0081)	(0.0080)	(0.0080)
	Retention secondary	0.0804***	0.0784***	0.0824***
		(0.0107)	(0.0110)	(0.0108)
	Retention secondary* percentage of retained students			-0.0013*
				(0.0006)
Variance components				
	Intercept U_0	0.0082***	0.0072***	0.0072***
	Gender U_1	0.0022***	0.0020***	0.0020***
	Age U_2	0.0004**	0.0004**	0.0004**
	Ethnicity U_3	0.0058*	0.0057*	0.0056*
	SES U_4	0.0001	0.0001	0.0001
	Vocational track U_5	0.0022**	0.0021**	0.0021**
	Prior achievement U_6	0.0001**	0.0001**	0.0001**
	Retention primary U_7	0.0009	0.0009	0.0009
	Retention secondary U_8	0.0035***	0.0038***	0.0039***

Note: Gamma coefficients are presented, with the standard errors appearing in parentheses, and the variance components. * $p < .05$; ** $p < .01$; *** $p < .001$

School retention composition and school misconduct

The second research question, whether school composition in terms of proportion of retained students influences school deviancy, is answered by model 2 (see Table 2). Controlling for retention in both primary and secondary education, we found a small but significant positive association between the percentage of retained students at school and school misconduct (γ=0.0018; p<0.001). This meant that both retained and promoted students in schools where proportionally more students had been retained in secondary education, had a higher likelihood of misbehaving. All else being equal, for two schools differing by one percentage point in their proportion of retained students, the misconduct rate of the one with the higher percentage was 100.18% (=100*Exp(0.0018)) that of the school with the lower percentage. This seemed like a very small difference. However, when we compared the misconduct rate of the school with the lowest percentage of retained students (0%), which was - all else being equal - 32.2817 (=Exp(3.4745)), to that of the school with the highest percentage (58%), which was 35.8341 (=Exp(3.4745+58*0.0018)), school retention composition did seem to matter.

The moderating role of schools' retention composition

The third research question, whether the association between grade retention and misconduct varies according to a school's retention composition, is answered in model 3 (see Table 2). First, the variance component of secondary school retention was significant in model 2 (U_8=0.0038, p<0.001), meaning that the slope of secondary school retention varied significantly between schools. Thus the introduction of cross-level interaction terms was warranted. The cross-level interaction effect between secondary school retention and the percentage of students retained was significantly negative (see model 3; γ=-0.0013; p<0.05). This meant that the deviance-yielding effect of secondary school retention on misconduct was diminished in schools where more students had been retained. All else being equal, for two schools differing by one percentage point in their percentage of retained students, the effect of secondary school retention on misconduct in the school with the higher percentage was 0.0013 less than that in the school with the lower percentage. Again, this seemed a very small difference. Yet if we compared the difference in misconduct rates between retained and promoted students in the school with the lowest percentage of retained students in our sample (0%) with those of the school with the highest (58%), the percentage of retained students at school did matter. In the school with the lowest percentage, retained students were 8.6% (=Exp(0.0824)) more likely to misbehave than non-retained students. In that with the highest percentage, retained students were 0.7% (=Exp(0.0824+58*(-0.0013)) more likely to break the rules than the non-retained.

DISCUSSION

The purpose of this study was to account for the shortcomings of research linking grade retention to students' disruptive behavior. First, in contrast to past research on this issue (see Jimerson & Ferguson, 2007), we considered retained students' likelihood of school deviancy in adolescence. As expected, our findings suggest that being retained is associated with

school-disruptive behavior. The results point to a rather strong deviance-yielding effect of retention during secondary education, so that our results seem to endorse the view that retained students are likely to feel relatively deprived in adolescence (Finlayson, 1977; Wu et al., 2010). However this study addresses the importance of distinguishing retention at different educational levels: since retention in primary school is associated with less disruptive behavior during adolescence, we seem to find no evidence that students retained in primary education feel relatively deprived in adolescence. This coincides with earlier research that suggested that the stigma of early retention may "wash away" over the years (Hong & Yu, 2008, p. 418). This positive image of primary grade retention, however, contradicts previous studies, most of which conclude that the practice should be avoided at the primary school level (e.g., Bonvin et al., 2008). In one of the few studies on outcomes in adolescence, Jimerson & Ferguson (2007) found that students retained during the early years of primary education display higher rates of aggression than those promoted. Our results, however, endorse the view expressed by the early intervention theory: early grade retention in primary education may make up for students' immaturity (e.g. poor behavioral regulation, poor concentration in class, etc.) (Hong & Yu, 2008; Wu et al., 2010). Immaturity is frequently given as a rationale for primary grade retention (Bowman, 2005), with an extra year of physical maturation possibly giving retained students the chance to conform better to classroom expectations than immature students who are promoted. Wu and colleagues (2010) find that retention in the early grades of primary education does lead to higher behavioral engagement of retained students. However, their data only extended to the end of primary education, and Wu and colleagues (2010) expressed doubts that these positive short-term effects of grade retention would last over the longer term. Our findings suggest that this effect may endure through adolescence, although we should point out that the positive effect seems to be quite small.

Furthermore, the results of this study underline the importance of accounting for the multilevel nature of the school context in assessing grade retention effects on students' disruptive behavior. We find students in schools where proportionally more have been retained to have a small, but significantly higher chance of breaking the rules than those in schools where fewer have been. It is noteworthy that these findings apply both to retained and promoted students. These results on students' school-disruptive behavior contradict the findings of Hong and Raudenbush (2006) with respect to students' achievement, who found that school retention composition did not affect students' cognitive growth. However, while Hong and Raudenbush (2006) used a dichotomous measure to subdivide schools into low-retention and high-retention ones, the substantial variation between schools in the Flemish educational context enables us to use a more refined, continuous measure to assess compositional effects (see also Hong & Raudenbush, 2006, p. 909). If the percentage of retained students in school can be seen as an indicator of a school's retention policy, as other authors have suggested (see Hong & Raudenbush, 2006; Shepard & Smith, 1988), we can state that our findings do not endorse the widespread use of grade retention in secondary schools.

Our results do show, however, that a school's retention composition may moderate the effect of grade retention on rule-breaking in secondary education. Here we find statistical evidence against the stable unit treatment value assumption (see Rubin, 1986), since the

outcome of retention for individual students seems to depend on the interventions other students in the same school receive. In line with social comparison theory (Festinger, 1954), students retained in low-retention schools are more likely to break rules than those retained in high-retention ones. This suggests that retention composition does indeed shape students' frame of reference for social comparison. When fewer fellow students have been retained, it appears that those who are retained are more likely to feel relatively deprived, which may result in disruptive behavior (see e.g., Rosenberg et al., 1989). However, when a larger number have been retained, social comparison is less likely to result in relative deprivation, making deviance less likely. The visibility of the other group therefore appears to be an important factor in the social comparison process (see e.g., Richer, 1976), so that for the retained, it may be more stigmatizing to attend low-retention schools than high-retention ones.

However, this moderation effect of secondary schools' retention composition should not be understood as encouraging the continuation of grade retention, and especially not in secondary education. All the students in our study retained during secondary education had a higher likelihood of misconduct, even those in schools where grade retention was practiced the most. Thus in regard to the secondary school context, we do not find evidence that grade retention has beneficial effects. This agrees with the majority of past research on its effectiveness (e.g., Holmes, 1989; Jimerson, 2001). Notably, no study has yet uncovered advantages of retention during adolescent years (Jimerson & Ferguson, 2007). Furthermore, if a practice with negative side-effects is applied more readily to some groups of students than to others, it may generate social inequality in schools. Previous research has shown that not all poor performers are even likely to be retained. For example, some teacher attitudes, such as their preference for retention (Bonvin et al., 2008) and some socio-demographic student characteristics, such as students' gender and socioeconomic background (Jimerson et al., 1997; Karweit, 1999; Meisels & Liaw, 1993), have been shown to affect their likelihood of grade retention. If retention is indeed an ineffective and discriminatory practice, it cannot be continued. In secondary education, it seems appropriate to search for other strategies to remedy poor educational performance. Other programs have been suggested, including summer schools, increased positive parental involvement, remedial activities during and after school hours, individualized educational programs, and so forth (for an extended discussion see Jimerson, 2001; Mccoy & Reynolds, 1999, p. 295). However, our results do favor early intervention theory (see also Hong & Yu, 2008): retention at an early age may make up for students' immaturity, which in the long run may ameliorate their behavior at school. Hence, contrary to many studies in the past (e.g., Bonvin et al., 2008; Jimerson et al., 2002), we find no evidence for the ineffectiveness of grade retention in primary education.

An important limitation of our study is that, contrary to most previous ones (e.g., Jimerson & Ferguson, 2007; Pagani et al., 2001), it does not use a longitudinal design. As a result, we cannot be sure about the direction of the effects. It may well be that school misconduct influences chances for grade retention, rather than the other way around (see e.g., Jimerson et al., 2002, p. 453). For example, disruptive students tend to have poorer academic achievement (see Roeser & Eccles, 1998), which makes them susceptible to be retained (Jimerson & Ferguson, 2007). Previous longitudinal research, however, shows that the relationship between grade retention and school misbehavior is likely to be bidirectional: disruptive students do have a larger chance of being retained, but retained students are also

more likely to be disruptive at a later point in time (Pagani et al., 2001). Moreover, this research confirms that the direction from grade retention to deviancy is dominant (Pagani et al., 2001). Aside from this, it is not very likely that primary school retention is caused by deviancy in adolescence, which by definition occurs at a later point in time. Furthermore, using longitudinal designs invokes a number of problems. As noted by Jimerson and Ferguson (2007, p. 332), longitudinal designs that span the beginning of primary education through to adolescence usually have to deal with attrition, resulting in small sample sizes which limit the generalizability of the results. Indeed, most retention studies in the past have been limited by small sample sizes and unrepresentative data (Ehmke et al., 2010; Lorence & Dworkin, 2006). Here we have accounted for this well-known critique of retention research by using an extensive dataset which is representative of the Flemish situation. Moreover, we account for another well-established critique of retention research by demonstrating that the multilevel nature of the school context should be taken into account in determining grade retention effects (Hong & Raudenbush, 2005). However, we do propose that future longitudinal research tries to replicate our findings, to account for the possible bidirectionality of misconduct and grade retention.

CONCLUSION

This study represents one of the few to investigate the effect of grade retention on students' school-disruptive behavior in adolescence. It is unique in addressing multilevel issues in this line of research. First, it has shown that it is important to distinguish grade retention at different educational levels. While we find evidence that primary school retention may be associated with less misconduct in adolescence, we establish that secondary school retention may give rise to deviance in adolescence. Moreover, we address the important role of schools' retention composition, finding that students attending schools with more retainees are more likely to be deviant, although this composition does moderate negative retention effects. Together with previous literature on the effectiveness of grade retention, we advocate the abandonment of this intervention, especially at the secondary level.

REFERENCES

Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (1994). *On the success of failure. A reassessment of the effects of retention in the primary grades.* Cambridge: University Press.

Barone, C., Weissberg, R. P., Kasprow, W. J., Voyce, C. K., Arthur, M. W., & Shriver, T. P. (1995). Involvement in multiple problem behaviors of young urban adolescents. *The Journal of Primary Prevention, 15,* 261-283.

Bonvin, P., Bless, G., & Schuepbach, M. (2008). Grade retention: decision-making and effects on learning as well as social and emotional development. *School Effectiveness and School Improvement, 19,* 1-19.

Bowman, L. J. (2005). Grade retention: Is it a help or hindrance to student academic success? *Preventing School Failure, 49,* 42-46.

Byrd, R. S., Weitzman, M., & Auinger, P. (1997). Increased behavior problems associated with delayed school entry and delayed school progress. *Pediatrics, 100,* 654-661.

Crosnoe, R., Erickson, K. G., & Dornbusch, S. M. (2002). Protective functions of family relationships and school factors on the deviant behavior of adolescent boys and girls. Reducing the impact of risky friendships. *Youth & Society, 33,* 515-544.

Demanet, J. & Van Houtte, M. (2011). Social-ethnic school composition and school misconduct: Does sense of futility clarify the picture? *Sociological Spectrum, 31,* 224-256.
Demanet, J. & Van Houtte, M. (2012). Teachers' attitudes and students' opposition. School misconduct as a reaction to teachers' diminished effort and affect. *Teaching and Teacher Education, 28,* 860-869.
Ehmke, T., Drechsel, B., & Carstensen, C. H. (2010). Effects of grade retention on achievement and self-concept in science and mathematics. *Studies in Educational Evaluation, 36,* 27-35.
Eitle, D. & Eitle, T. M. (2003). Segregation and school violence. *Social Forces, 82,* 589-616.
Festinger, L. A. (1954). A theory of social comparison processes. *Human Relations, 7,* 117-140.
Finlayson, H. J. (1977). Non-promotion and self-concept development. *Phi Delta Kappan, 59,* 205-206.
Frost, M. B. (2007). Texas students' college expectations: Does high school racial composition matter? *Sociology of Education, 80,* 43-65.
Gottfredson, D. C., Fink, C. M., & Graham, N. (1994). Grade retention and problem behavior. *American Educational Research Journal, 31,* 761-784.
Holmes, C. T. (1989). Grade level retention effects: a meta-analysis of research studies. In L.A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 16-33). London: Falmer Press.
Hong, G. & Raudenbush, S. W. (2005). Effects of Kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis, 27,* 205-224.
Hong, G. & Raudenbush, S. W. (2006). Evaluating Kindergarten retention policy: A case study of causal inference for multilevel observational data. *Journal of the American Statistical Association, 101,* 901-910.
Hong, G. & Yu, B. (2008). Effects of Kindergarten retention on children's social-emotional development: An application of propensity score method to multivariate, multilevel data. *Developmental Psychology, 44,* 407-421.
Huisman, M. (1999). Imputation of missing item responses: some simple techniques. In M.Huisman (Ed.), *Item Nonresponse: Occurence, Causes, and Imputation of Missing Answers to Test Items.* (pp. 91-119). Leiden: DSWO Press, Leiden University.
Jackson, G. B. (1975). Research evidence on effects of grade retention. *Review of Educational Research, 45,* 613-635.
Jimerson, S., Carlson, E., Rotert, M., Egeland, B., & Sroufe, L. A. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology, 35,* 3-25.
Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review, 30,* 420-437.
Jimerson, S. R., Anderson, G. E., & Whipple, A. D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools, 39,* 441-457.
Jimerson, S. R. & Ferguson, P. (2007). A longitudinal study of grade retention: Academic and behavioral outcomes of retained students through adolescence. *School Psychology Quarterly, 22,* 314-339.
Juchtmans, G., Belfi, B., De Fraine, B., Goos, M., Knipprath, H., Vandenbroucke, A. et al. (2011). *Samen tot aan de meet. Alternatieven voor zittenblijven. [Together until the end. Alternatives for retention]*. Antwerpen: Garant.

Karweit, N. L. (1999). *Grade retention: Prevalence, timing, and effects.* Baltimore: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk (Report No. 33).
Lee, V. E. & Burkam, D. T. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal, 40,* 353-393.
Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Thousand Oaks, CA: Sage.
Lorence, J. (2006). Retention and academic achievement research revisited from a United States perspective. *International Educational Journal, 7,* 731-777.
Lorence, J. & Dworkin, A. G. (2006). Elementary grade retention in Texas and reading achievement among racial groups: 1994-2002. *Review of Policy Research, 23,* 999-1033.
Mccoy, A. R. & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37,* 273-298.
Meisels, S. J. & Liaw, F. R. (1993). Failure in grade. Do retained students catch up? *Journal of Educational Research, 87,* 69-77.
Opdenakker, M-C., & Van Damme, J. (1997). Centreren in multilevelanalyse: implicaties van twee centreringsmethoden voor het bestuderen van schooleffectiviteit [Centering in multilevel analysis: Implications for two methods of centering for examining school effectiveness]. *Tijdschrift voor Onderwijsresearch, 22,* 264-290.
Pagani, L., Tremblay, R. E., Vitaro, F., Boulerice, B., & McDuff, P. (2001). Effects of grade retention on academic performance and behavioral development. *Development and Psychopathology, 13,* 297-315.
Peterson, S. E., Degracie, J. S., & Ayabe, C. R. (1987). A longitudinal study of the effects of retention/promotion on academic achievement. *American Educational Research Journal, 24,* 107-118.
Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. London: Sage Publications.
Richer, S. (1976). Reference group theory and ability grouping: A convergence of sociological theory and educational research. *Sociology of Education, 49,* 65-71.
Roeser, R. W. & Eccles, J. S. (1998). Adolescents' perceptions of middle school: Relation to longitudinal changes in academic and psychological adjustment. *Journal of Research on Adolescence, 8,* 123-158.
Rosenberg, M., Schooler, C., & Schoenbach, C. (1989). Self-esteem and adolescent problems: Modeling reciprocal effects. *American Sociological Review, 54,* 1004-1018.
Rubin, D. B. (1986). Comment: Which ifs have causal answers. *Journal of the American Statistical Association, 81,* 961-962.
Shepard, L. A. & Smith, M. L. (1988). Escalating academic demand in Kindergarten: Counterproductive policies. *The Elementary School Journal, 89,* 135-145.
Shepard, L. A. & Smith, M. L. (1990). Synthesis of research on grade retention. *Educational Leadership, 47,* 84-88.
Stearns, E., Moller, S., Blau, J., & Potochnick, S. (2007). Staying back and dropping out: The relationship between grade retention and school dropout. *Sociology of Education, 80,* 210-240.
Stevens, P. A. J. (2007). Exploring the importance of teachers' institutional structure on the development of teachers' standards of assessment in Belgium. *Sociology of Education, 80,* 314-329.
Stewart, E. A. (2003). School social bonds, school climate, and school misbehavior: A multilevel analysis. *Justice Quarterly, 20,* 575-604.

Stretesky, P. B. & Hogan, M. J. (2005). Segregation and school disorder. *Social Science Journal, 42,* 405-420.

Trautwein, U., Ludtke, O., Marsh, H. W., Koller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology, 98,* 788-806.

Tygart, C. E. (1988). Strain theory and public school vandalism. Academic tracking, school social status, and students' academic achievement. *Youth and Society, 20,* 106-118.

Van Houtte, M., Demanet, J., & Stevens, P. A. J. (2012). Self-esteem of academic and vocational students: Does within-school tracking sharpen the difference? *Acta Sociologica, 55,* 73-89.

Van Houtte, M. & Stevens, P. A. J. (2008). Sense of futility: The missing link between track position and self-reported school misconduct. *Youth and Society, 40,* 245-264.

Van Houtte, M., Stevens, P.A.J., Sels, A., Soens, K. & Van Rossem, R. (2005). De invloed van structurele en compositorische schoolkenmerken op prestaties en welbevinden van leerlingen in het secundair onderwijs: een verklaring via cultuur. [The influence of structural and compositional school features on achievement and well-being of students in secondary education: an explanation through culture]. First research report (not published). Ghent: Universiteit Gent, vakgroep sociologie, onderzoeksgroep jeugd, educatie en gender.

Van Petegem, P. & Schuermans, G. (2005). Zittenblijven in Vlaanderen. De relatie tussen dubbelen in het secundair onderwijs en de schoolloopbaan in het hoger onderwijs [Retention in Flanders: The relation between secondary school retention and the educational career in higher education]. *Impuls, 36,* 3-12.

Wu, W., West, S. G., & Hughes, J. N. (2010). Effect of grade retention in first grade on psychosocial outcomes. *Journal of Educational Psychology, 102,* 135-152.

APPENDIX: The School Misconduct Scale (inspired by Stewart, 2003, p. 602-604)

How often have you:

1. been late for school
2. skipped lessons
3. skipped school all day
4. cheated on tests
5. copied someone's homework
6. not made your homework
7. fought at school
8. stolen at school
9. committed vandalism at school
10. smoked at school
11. drunk alcohol during school hours
12. done drugs during school hours
13. talked back at teachers
14. broke the school rules
15. had to do impositions
16. been sent to detention
17. been suspended for one or more days